CALIFORNIA STATE UNIVERSITY, NORTHRIDGE

WORKING TITLE

A thesis submitted in partial fulfillment of the requirements for
the degree of Master of Science in Computer Science

By

Jeffrey Limbacher

August 2015

The thesis of Jeffrey Limbacher is approved:

Ali Zakeri , Ph.D. Date

Vladislav Panferov , Ph.D. Date

Alexander Alekseenko , Ph.D., Chair Date

California State University, Northridge

Acknowledgements

tbd

Table of Contents

ABSTRACT

WORKING TITLE

By

Jeffrey Limbacher

Master of Science in Computer Science

tbd

Chapter 1

Introduction

Chapter 2

The Boltzmann Equation

The kinematic theory of gases treats gases as composed of a large number of individual molecules that for large periods of time flow freely. As these particles move freely through space, they collide with each other. Collisions of these particles are what drive the evolution of the gas towards equilibrium.

2.1 Binary Collisions of Particles

This section considers the properties of two particles on a collision path with each other as illustrated in Figure 2.1. In all the work that follows, is is assumed that the molecules undergo elastic hard sphere collisions.

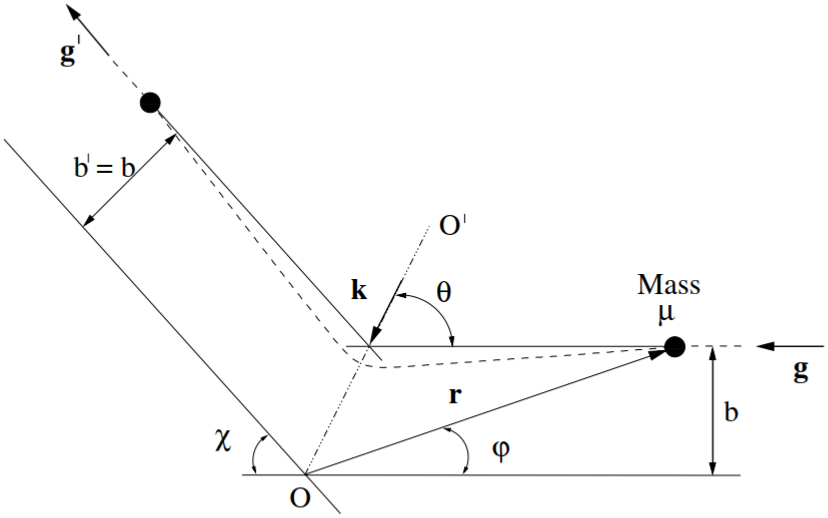

Figure 2.1: Kremer (2010, p. 27), Fig 1.6

Denote the pre- and post-collisional asymptotic velocities by $\vec{v}$, $\vec{v_1}$ and $\vec{v}'$, $\vec{v}_1'$ respectively. Define the relative pre- and post-collisional velocities, respectively, by

$$\vec{g} = v_1 - v\,, \quad \vec{g}' = \vec{v}' - \vec{v}_1'\,.$$

b denotes the offset of the centers of the molcules orthogonal to $\vec{g}$. ε denotes the azimuthal angle between the two particles.

By conservation of momentum, we have that

$$m\vec{v} + m\vec{v}_1 = m\vec{v}' + m\vec{v}_1' \,. \tag{2.1}$$

Equation 2.1 yields $|g'| = |g|$. In addition, due to the hard sphere assumption, the collision is considered to be perfectly elastic, giving

$$\frac{1}{2}m|\vec{v}| + \frac{1}{2}m|\vec{v}_1| = \frac{1}{2}m|\vec{v}'| + \frac{1}{2}m|\vec{v}_1'| \tag{2.2}$$

The apsidal vector, $\vec{k}$ given by

$$\vec{k} = \frac{\vec{g} - \vec{g}'}{|\vec{g} - \vec{g}'|} \,,$$

bisects the angle between asymptotic relative velocities. Using this vector, we can write a relationship between the pre- and post-collisional velocities by

$$\vec{v}_1' = \vec{v}_1 - \vec{k}(\vec{k} \cdot \vec{g}) \,, \quad \vec{v}' = \vec{v} + \vec{k}(\vec{k} \cdot \vec{g}) \,. \tag{2.3}$$

2.2 The Boltzmann Equation

We consider a gas enclosed in a volume. A single molecule of this gas can be described with having position $\vec{x}$ and velocity $\vec{v}$ at a time t. For a particular time, we can describe a molecule as being within at a single point, $(\vec{x}, \vec{v})$, in 6-dmensional space known as phase space. We define the distribution function of the gas as $f(t, \vec{x}, \vec{v})d\vec{x}\, d\vec{v}$ gives the number of particles within the range of $\vec{x} + d\vec{x}$ with velocities $\vec{v} + d\vec{v}$.

In 1872 Boltzmann [4] introduced the Boltzmann equation which describes the time evolution of the distribution f. In the absence of external forces and we ignore collisions of particles, then the Boltzmann equation takes the form

$$\frac{\partial}{\partial t} f(t, \vec{x}, \vec{v}) + \vec{v} \cdot \nabla_x f(t, \vec{x}, \vec{v}) = 0 \,. \tag{2.4}$$

However, when the effects of collisions cannot be neglected, the right hand side must be modified to take this into account. In this case, the Boltzmann equation takes the form of

$$\frac{\partial}{\partial t} f(t, \vec{x}, \vec{v}) + \vec{v} \cdot \nabla_x f(t, \vec{x}, \vec{v}) = I[f](t, \vec{x}, \vec{v}) \tag{2.5}$$

Where $I[f]$ is referred to as the collision operator. The explicit form of $I[f]$ depends on the properties of the gas. To describe it explicitly, we make several assumptions. First, the gas composed entirely of a single species of molecule. Second, we assume hard sphere collisions as described in section 2.1.

In order to explicitly write the collision operator, we must describe the collisions within the gas. Consider a particle with velocity $\vec{v}_1$ at point $\vec{x}$. This particle will collide with $f(t, \vec{x}, \vec{v}) d\vec{v}_1 g \Delta t b \, db \, d\varepsilon$ particles within the ranges of $\vec{v}_1$ and $\vec{v}_1 + d\vec{v}_1$. Integrating this by all all possible angles ($0 <= \varepsilon < 2\pi$), over all impact parameters, ($0 \leq b \leq b^*$), and over all velocities, for all particles within the volume element $\vec{x} d\vec{x}$, we get that there are

$$dt \int_{\mathbb{R}^3} \int_0^{b^*} \int_0^{2\pi} f(t, \vec{x}, \vec{v}_1) f(t, \vec{x}, \vec{v}) b |g| \, db \, d\varepsilon \, d\vec{v}_1 \tag{2.6}$$

total collisions within the volume element that annihilate points with velocity $\vec{v}$. Likewise, there are collisions that create points with velocity $\vec{v}$. Using the results of the lasts section, points with velocity $\vec{v}'$ and $\vec{v}_1'$, impact parameter $b' = b$, and azimuthal angle $\varepsilon' = \pi + \varepsilon$ will result in particles with post-collisional velocity of $\vec{v}$

and $\vec{v}_1$. The number of such collisions is

$$\begin{aligned} &dt\int_{\mathbb{R}^3}\int_0^{b^*}\int_0^{2\pi} f(t,\vec{x},\vec{v}_1')f(t,\vec{x},\vec{v}')b'|g'|\,db\,d\varepsilon'\,d\vec{v}_1 \\ =&dt\int_{\mathbb{R}^3}\int_0^{b^*}\int_0^{2\pi} f(t,\vec{x},\vec{v}_1')f(t,\vec{x},\vec{v}')b|g|\,db\,d\varepsilon\,d\vec{v}_1\,. \end{aligned} \tag{2.7}$$

Subtracting (2.6) from (2.7) gives the net number of particles enter or leave the volume element $\vec{v}+d\vec{v}$; that is

$$I[f](t,\vec{x},\vec{v}) = \int_{\mathbb{R}^3}\int_{\mathbb{R}^3}(f_1'f'-f_1f)|g|\,b\,db\,d\varepsilon\,d\vec{v}_1\,, \tag{2.8}$$

where,

$$f_1' \equiv f(t,\vec{x},\vec{v}_1') \quad f' \equiv f(t,\vec{x},\vec{v}') \quad f_1 \equiv f(t,\vec{x},\vec{v}_1 \quad f \equiv f(t,\vec{x},\vec{v})\,.$$

From here, we can substitute (2.8) into (2.5) to arrive to the explicit form of the Boltzmann equation,

$$\frac{\partial}{\partial t}f(t,\vec{x},\vec{v}) + \vec{v}\cdot\nabla_x f(t,\vec{x},\vec{v}) = \int_{\mathbb{R}^3}\int_{\mathbb{R}^3}(f_1'f'-f_1f)|g|\,b\,db\,d\varepsilon\,d\vec{v}_1\,. \tag{2.9}$$

The Boltzmann equation is a non-linear integro-differential equation. The right hand side is a five dimensional integral that must be evaluated at each point in 6-dimensional space.

2.3 Moments of the Distribution Function

A gas is usually described by its macroscopic states. Kinetic theory defines these macroscopic properties in terms of distribution function $f(t,\vec{x},\vec{v})$. The first five moments of the gas are defined below.

$$
\begin{aligned}
n(t,\vec{x}) &= \int_{\mathbb{R}^3} f(t,\vec{x},\vec{v})d\vec{v} && \text{- number density} && (2.10)\\
\bar{v}_i(t,\vec{x}) &= \frac{1}{n(t,\vec{x})}\int_{\mathbb{R}^3} mv_i f(t,\vec{x},\vec{v})d\vec{v} && \text{- bulk velocity} && (2.11)\\
T(t,\vec{x}) &= \frac{1}{3Rn(t,\vec{x})}\int_{\mathbb{R}^3} mC^2 f(t,\vec{x},\vec{v})d\vec{v} && \text{- temperature} && (2.12)
\end{aligned}
$$

where $C_i = v_i - \bar{v}_i$, $C^2 = C_1^2 + C_2^2 + C_3^2$, and R is the specific gas constant. The number density $n(t,\vec{x})$ denote the number of particles contained in our distribution. $\vec{\bar{v}}(t,\vec{x})$ denotes that average velocity of particles within the gas. The temperature denotes the deviation form the average.

An important property of the collision integral is that the first five moments are conservative (see [7]), that is

$$
\begin{aligned}
\int_{\mathbb{R}^3} I[f](t,\vec{x},\vec{v})d\vec{v} = 0\,,\\
\int_{\mathbb{R}^3} v_i I[f](t,\vec{x},\vec{v})d\vec{v} = 0\,,\\
\int_{\mathbb{R}^3} C^2 I[f](t,\vec{x},\vec{v})d\vec{v} = 0\,.
\end{aligned}
\tag{2.13}
$$

2.4 The Maxwellian Distribution

If the gas is free from external influence, then the gas will approach an equilibrium. In this equilibrium, the gas distribution takes a specific shape known as the Maxwellian distribution given below.

$$
f_M(\vec{v}, n, \vec{\bar{v}}, T) = \frac{1}{\sqrt{2\pi RT}^3}\exp\left(-\frac{|\vec{v}-\vec{\bar{v}}|^2}{2RT}\right) \tag{2.14}
$$

Note that the exact shape of the Maxwellian distribution depends on the macroscopic moments of the gas distribution, n, $\vec{\bar{v}}$, and T given by equations 2.10, 2.11, 2.12

respectively. This is illustrated in Figure 2.2. The distribution is centered around $\vec{\bar{v}}$. The temperature, T, controls the width of the distribution. n determines the area under the curve.

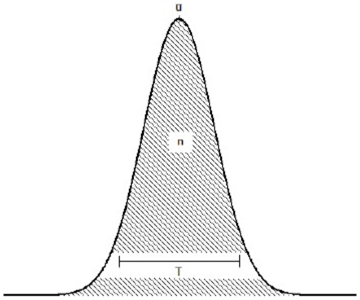

Figure 2.2: Stolen picture need better caption

When the gas is in equilibrium, the difference between the number of particles that enter and leave a particular phase volume vanishes. In other words,

$$I[f_M](t,\vec{x},\vec{v}) = \int_{\mathbb{R}^3}\int_{\mathbb{R}^3} (f'_{M1}f'_M - f_{M1}f_M)g\, b\, db\, d\varepsilon\, d\vec{v}_1 \;\; = 0$$

2.5 Dimensionless Reduction

Gas dynamic constants can vary greatly in scale. This can cause an accumulation of round-off error using floating point arithmetic when performing a large number of computations. In order to reduce the error that can be introduced by the varying scales, a dimensionless reduction is performed on the constants and equations. The dimensionless reduction aims to reduce the scale of all the variables to roughly of order one to minimize the round-off error. Note that the dimensionless reduction process can never eliminate round-off error since it is inherit to floating point arithmetic. The techniques used for dimensionless reduction vary from problem to problem. This thesis adopts the convention borrowed from Chapter 3 of [][][][].

Chapter 3

Discontinuous Galerkin Discretization in the Velocity Variable

Discontinuous Galerkin methods is a method of discretizing equations. The following section describes the DG formulation found in [1], [3].

3.1 DG Discretization in Velocity Space

We denote the points in the velocity space as $\vec{v} = (u, v, w)$. The velocity space is reduced to a rectangular parallelpiped $K = [u_L, u_R] \times [v_L, v_R] \times [w_L, w_R]$. It is assumed that outside the parallelpiped the contribution of the function to the first few moments is negligible. Depending on the parallelpiped, this will not result in large errors in terms of conservation of density and temperature.

We partition K into N smaller rectangular parallelpipeds $K_j = [u^j_L, u^j_R] \times [v^j_L, v^j_R] \times [w^j_L, w^j_R]$. Each K_j will contain a set of basis function, ϕ^i_j, $i = 1, \dots, s$ as described. We introduce nodes of Gauss quadratures of order s_u, s_v, and s_w on each of the intervals $[u^j_L, u^j_R]$, $[v^j_L, v^j_R]$, and $[w^j_L, w^j_R]$. The nodes are denoted as $\kappa^u_{p;j}$, $\dots$, $p = 1, s_u$, $\kappa^v_{q;j}$, $\dots$, $q = 1, s_v$, and $\kappa^w_{r;j}$, $\dots$, $r = 1, s_w$. From the nodes, the one-dimensional Lagrange basis functions are defined:

$$\phi^u_{l;j}(u) = \prod_{\substack{p=1,s_u \\ p\neq l}} \frac{\kappa^u_{p;j} - u}{\kappa^u_{p;j} - \kappa^u_{l;j}}, \quad \phi^v_{m;j}(v) = \prod_{\substack{q=1,s_v \\ q\neq m}} \frac{\kappa^v_{q;j} - v}{\kappa^v_{q;j} - \kappa^v_{m;j}}, \quad \phi^w_{n;j}(w) = \prod_{\substack{r=1,s_w \\ r\neq n}} \frac{\kappa^w_{r;j} - w}{\kappa^w_{r;j} - \kappa^w_{n;j}}. \tag{3.1}$$

The three-dimensional basis function is defined as

$$\phi_{i;j}(\vec{v}) = \phi^u_{l;j}(u)\phi^v_{m;j}(v)\phi^w_{n;j}(w) \tag{3.2}$$

where $l = 1, \dots, s_u$, $m = 1, \dots, s_v$, $n = 1, \dots, s_w$ and i is the index that runs through all possible combinations of l, n, and m, and is computed as $i = ((l-1)s_v) + (m-1)s_w) + n$. A useful property of the basis functions (3.2) is that they vanish on all nodes except one. In addition, the quadrature nodes used are exact on polynomials of degree at most $2s_u - 1$, $2s_v - 1$, and $2s_w - 1$. In addition, the following lemma holds,

Lemma 1 *(see also [2, 5]) The following identities hold for basis functions $\phi_{i;j}(\vec{v})$:*

$$\int_{K_j} \phi_{p;j}(\vec{v})\phi_{q;j}(\vec{v})\, d\vec{v} = \frac{\omega_p \Delta\vec{v}^j}{8}\delta_{pq} \qquad \text{and} \qquad \int_{K_j} \vec{v}\phi_{p;j}(\vec{v})\phi_{q;j}(\vec{v})\, d\vec{v} = \frac{\omega_p \Delta\vec{v}^j}{8}\vec{v}_{p;j}\delta_{pq}, \tag{3.3}$$

where indices l, n, and m of one dimensional basis functions correspond to the three-dimensional basis functions $\phi_{p;j}(\vec{v}) = \phi^u_{l;j}(u)\phi^v_{m;j}(v)\phi^w_{n;j}(w)$, and the vector $\vec{v}_{p;j} = (\kappa^u_{l;j}, \kappa^v_{m;j}, \kappa^w_{n;j})$.

3.2 Nodal-DG Velocity Discretization of the Boltzmann Equation

We assume that on each K_j, the solution to the Boltzmann equation is sought of the form

$$f(t,\vec{x},\vec{v})|_{K_j} = \sum_{i=1,s} f_{i;j}(t,\vec{x})\phi_{i;j}(\vec{v})\,. \tag{3.4}$$

We substitute equation 3.4 into 2.5, multiply the result by test basis function, integrate over K_j, and apply identity (3.3) to arrive to

$$\partial_t f_{i;j}(t,\vec{x}) + \vec{v}_{i;j}\cdot\nabla_x f_{i;j}(t,\vec{x}) = \frac{8}{\omega_i \Delta\vec{v}^j} I_{\phi_{i;j}}\,, \tag{3.5}$$

where $I_{\phi_{i;j}}$ is the projection of the collision operator on the basis function $\phi_{i;j}(\vec{v})$:

$$I_{\phi_{i;j}} = \int_{K_j} \phi_{i;j}(\vec{v}) I[f](t,\vec{x},\vec{v})\, d\vec{v}\,. \tag{3.6}$$

3.3 Reformulation of the Galerkin Projection of the Collision Operator

Similarly to [1, 2, 8], we rewrite the DG projection of the collision operator $I_{\phi_{i;j}}$ in the form of a bilinear integral operator with a time-independent kernel. The principles of kinetic theory suggest that changes to $f(t,\vec{x},\vec{v})$ with respect to $\vec{x}$ at the distance of a few b^* are negligible, see e.g., [10]. Specifically, using the well-known identities (see, e.g., [6], Section 2.4), and applying the first principles assumption, we have

$$\begin{aligned} I_{\phi_{i;j}} &= \int_{\mathbb{R}^3}\int_{\mathbb{R}^3} f(t,\vec{x},\vec{v})f(t,\vec{x},\vec{v}_1)\int_{\mathbb{S}^2}(\phi_{i;j}(\vec{v}') - \phi_{i;j}(\vec{v}))b_\alpha(\theta)|g|^\alpha \, d\sigma\, d\vec{v}_1\, d\vec{v} \\ &= \int_{\mathbb{R}^3}\int_{\mathbb{R}^3} f(t,\vec{x},\vec{v})f(t,\vec{x},\vec{v}_1)A(\vec{v},\vec{v}_1;\phi_{i;j})d\vec{v}_1\, d\vec{v}\,, \end{aligned} \tag{3.7}$$

where

$$A(\vec{v},\vec{v}_1;\phi_{i;j}) = |g|^\alpha \int_{\mathbb{S}^2}(\phi_{i;j}(\vec{v}') - \phi_{i;j}(\vec{v}))b_\alpha(\theta)\, d\sigma\,. \tag{3.8}$$

The kernel $A(\vec{v},\vec{v}_1;\phi_{i;j})$ is independent of time and can be pre-computed. In [2] properties of a kernel closely related to $A(\vec{v},\vec{v}_1;\phi_{i;j})$ are considered. In particular, due to the local support of $\phi_{i;j}(\vec{v})$, it is anticipated that kernel $A(\vec{v},\vec{v}_1;\phi_{i;j})$ will have only $O(M^5)$ non-zero components for each $\phi_{i;j}(\vec{v})$, where M is the number of discrete velocity points in each velocity dimension. As a result, evaluation of (3.7) will require $O(M^8)$ operations for each spatial point. This number of evaluations is very high. However, as we will show later, it can be reduced to $O(M^6)$ operations using symmetries of $A(\vec{v},\vec{v}_1;\phi_{i;j})$, the convolution form of (3.7) and the Fourier transform.

We remark that in many numerical re-formulations of the Boltzmann equation, the collision operator is separated into the gain and loss terms. This separation can be performed in (3.7),

$$\begin{aligned} I_{\phi_{i;j}} &= \int_{\mathbb{R}^3}\int_{\mathbb{R}^3} f(t,\vec{x},\vec{v})f(t,\vec{x},\vec{v}_1)\int_{\mathbb{S}^2}(\phi_{i;j}(\vec{v}') - \phi_{i;j}(\vec{v}))b_\alpha(\theta)|g|^\alpha \, d\sigma \, d\vec{v}_1 \, d\vec{v} \\ &= \int_{\mathbb{R}^3}\int_{\mathbb{R}^3} f(t,\vec{x},\vec{v})f(t,\vec{x},\vec{v}_1)(\int_{\mathbb{S}^2}\phi_{i;j}(\vec{v}')b_\alpha(\theta)|g|^\alpha d\sigma \; - |g|^\alpha \int_{\mathbb{S}^2}\phi_{i;j}(\vec{v})b_\alpha(\theta)d\sigma \;)d\vec{v}_1 \, d\vec{v} \\ &= \int_{\mathbb{R}^3}\int_{\mathbb{R}^3} f(t,\vec{x},\vec{v})f(t,\vec{x},\vec{v}_1)(A^+(\vec{v},\vec{v}_1;\phi_{i;j}) - |g|^\alpha \sigma_T)d\vec{v}_1 \, d\vec{v} \end{aligned} \tag{3.9}$$

where

$$A^+(\vec{v},\vec{v}_1;\phi_{i;j}) = \int_{\mathbb{S}^2}\phi_{i;j}(\vec{v}')b_\alpha(\theta)|g|^\alpha d\sigma\,, \quad \sigma_T = \int_{\mathbb{S}^2}\phi_{i;j}(\vec{v})b_\alpha(\theta)d\sigma\,. \tag{3.10}$$

$A^+(\vec{v},\vec{v}_1;\phi_{i;j})$ has similar properties to that of $A(\vec{v},\vec{v}_1;\phi_{i;j})$, but $A^+(\vec{v},\vec{v}_1;\phi_{i;j})$ has certain properties that are better than that of $A(\vec{v},\vec{v}_1;\phi_{i;j})$ for a Fourier transform. $A(\vec{v},\vec{v}_1;\phi_{i;j})$ is grows linearly to infinity in some direction whereas $A^+(\vec{v},\vec{v}_1;\phi_{i;j})$ does not. It can then be argued that $A^+(\vec{v},\vec{v}_1;\phi_{i;j})$ is better suited for a Fourier transform than $A(\vec{v},\vec{v}_1;\phi_{i;j})$. The algorithms in this paper also have straightforward extensions to the the split formulation of (3.7). However, in practice, the split formulation of (3.7) had significant errors in conservation.

3.4 Properties of the Kernel $A(\vec{v},\vec{v}_1;\phi_{i;j})$

Lemma 2 *Let operator $A(\vec{v},\vec{v}_1;\phi_{i;j})$ be defined by (3.8). Then $\forall \xi \in \mathbb{R}^3$*

$$A(\vec{v}+\vec{\xi},\vec{v}_1+\vec{\xi};\phi_{i;j}(\vec{v}-\vec{\xi})) = A(\vec{v},\vec{v}_1;\phi_{i;j})\,.$$

Proof. Consider $A(\vec{v}+\vec{\xi},\vec{v}_1+\vec{\xi};\phi_{i;j}(\vec{v}-\vec{\xi}))$. We clarify that these notations mean that particle velocities $\vec{v}$ and $\vec{v}_1$ in (3.8) are replaced with $\vec{v}+\vec{\xi}$ and $\vec{v}_1+\vec{\xi}$ correspondingly and that basis function $\phi_{i;j}(\vec{v})$ is replaced with a "shifted" function $\phi_{i;j}(\vec{v}-\vec{\xi})$. We notice that the relative speed of the molecules with velocities $\vec{v}+\vec{\xi}$ and $\vec{v}_1+\vec{\xi}$ is still

$\vec{g} = \vec{v}+\vec{\xi}-(\vec{v}_1+\vec{\xi}_1) = \vec{v}-\vec{v}_1$. The post-collision velocities for the pair of particles will be $\vec{v}'+\vec{\xi}$ and $\vec{v}_1'+\vec{\xi}$, where $\vec{v}'$ and $\vec{v}_1'$ are given by (2.3). We notice, in particular, that choices of θ and ε in (2.3) are not affected by $\vec{\xi}$. The rest of the statement follows by a direct substitution:

$$A(\vec{v}+\vec{\xi},\vec{v}_1+\vec{\xi};\phi_{i;j}(\vec{v}-\vec{\xi})) = |g|^{\alpha}\int_{\mathbb{S}^2}\phi_{i;j}((\vec{v}'+\vec{\xi})-\vec{\xi})b_{\alpha}(\theta)\,d\sigma\ = |g|^{\alpha}\int_{\mathbb{S}^2}\phi_{i;j}(\vec{v}')b_{\alpha}(\theta)\,d\sigma$$

$$= A(\vec{v},\vec{v}_1;\phi_{i;j})\,.$$

■

We remark that Lemma 2 holds for all potentials of molecular interaction used in rarefied gas dynamics. This property was used in [2] to reduce the storage requirement on uniform partitions.

3.5 Rewriting the Collision Operator in the Form of a Convolution (Mostly Copy and Pasted from the Paper, needs rewrite?)

It was shown in [3] that the Galerkin projection of the collision operator can be reformulated in terms of a convolution. This work is recalled in this section. We select a partition cell K_c and designate this cell as a generating cell. Similarly, the basis functions $\phi_{i;c}(\vec{v})$ on K_c are designated as the generating basis functions. Basis functions $\phi_{i;j}(\vec{v})$ on other cells can be obtained using a shift in the velocity variable, namely $\phi_{i;j}(\vec{v}) = \phi_{i;c}(\vec{v}+\vec{\xi}_j)$ where $\vec{\xi}_j \in \mathbb{R}^3$ is the vector that connects the center of K_j to the center of K_c.

According to Lemma 2, operator $A(\vec{v},\vec{v}_1,\phi_{i;j})$ is invariant with respect to translations. Therefore

$$\begin{aligned} I_{\phi_{i;j}} &= \int_{\mathbb{R}^3}\int_{\mathbb{R}^3} f(t,\vec{x},\vec{v})f(t,\vec{x},\vec{v}_1)A(\vec{v}+\vec{\xi}_j,\vec{v}_1+\vec{\xi}_j;\phi_{i;j}(\vec{u}-\vec{\xi}_j))\,d\vec{v}_1 d\vec{v} \\ &= \int_{\mathbb{R}^3}\int_{\mathbb{R}^3} f(t,\vec{x},\vec{v})f(t,\vec{x},\vec{v}_1)A(\vec{v}+\vec{\xi^j},\vec{v}_1+\vec{\xi}_j;\phi_{i;c}(\vec{u}))\,d\vec{v}d\vec{v}_1\,. \end{aligned} \tag{3.11}$$

Performing the substitutions $\hat{\vec{v}} = \vec{v}+\vec{\xi}_j$ and $\hat{\vec{v}}_1 = \vec{v}_1+\vec{\xi}_j$ in (3.11), we have

$$I_{\phi_{i;j}} = \int_{\mathbb{R}^3}\int_{\mathbb{R}^3} f(t,\vec{x},\hat{\vec{v}}-\vec{\xi}_j)f(t,\vec{x},\hat{\vec{v}}_1-\vec{\xi}_j)A(\hat{\vec{v}},\hat{\vec{v}}_1;\phi_{i;c}(\vec{u}))\,d\hat{\vec{v}}d\hat{\vec{v}}_1\,.$$

We then introduce a bilinear convolution operator, $i=1,\dots,s$

$$I_i(\vec{\xi}) = \int_{\mathbb{R}^3}\int_{\mathbb{R}^3} f(t,\vec{x},\vec{v}-\vec{\xi})f(t,\vec{x},\vec{v}_1-\vec{\xi})A(\vec{v},\vec{v}_1;\phi_{i;c})\,d\vec{v}d\vec{v}_1\,, \tag{3.12}$$

and notice that $I_{\phi_{i;j}}$ can be obtained from (3.12) as $I_{\phi_{i;j}} = I_i(\vec{\xi}_j)$. In the following, we will refer to (3.12) as the convolution form of the Galerkin projection of the collision integral.

3.6 Discretization of the Collision Integral (COpy and pasted again)

In order to calculate (3.12), we replace the three-dimensional integrals with the Gauss quadratures associated with then nodal-DG discretization, (3.2). As is discussed above, we are only interested in computing convolution (3.12) at vectors $\vec{\xi}=\vec{\xi}_j$ that connect centers of the velocity cells K_j to the center of the velocity cell K_c, the support of $\phi_{i;c}(\vec{v})$. Since the same nodal points are used on all velocity cells, shifts $\vec{\xi}_j$ translate nodal points in one cell to nodal points in another cell. As a result, the quadrature sums to evaluate convolution (3.12) use values of the unknown $f(t,\vec{x},\vec{v})$ at the nodal points only. In fact, the shift in the velocity variable $\vec{v}_{i;l}-\vec{\xi}_j$ will correspond to a shift in the three dimensional index of the velocity cell which we will write formally as $l-j$, producing the velocity node $\vec{v}_{i;l-j}(\vec{v})$. The exact expression for the shift $l-j$ will be made clear later by considering the cell indices in each velocity dimension. The index i of the node within the cell is not affected in this process.

We can write the discrete form of (3.12) as

$$I_{i;j} := I_i(\vec{\xi}_j) = \sum_{i',i''=1}^{s} \sum_{j'=1}^{M^3} \sum_{j''=1}^{M^3} f_{i';j'-j} f_{i'';j''-j} A_{i',i'';j',j'';i} \tag{3.13}$$

where $f_{i';j'-j} = f(t,\vec{x},\vec{v}_{i';j'-j})$, $A_{i',i'';j',j'';i} = A(\vec{v}_{i';j'},\vec{v}_{i'';j''};\phi_{i;c})$ and the three dimensional indices i' and i'' run over the velocity nodes within a single velocity cell and indices j' and j'' run over all velocity cells. We note that for some index shifts $j'-j$, the resulting cells are outside of the velocity domain. In [2] the values outside of the domain were substituted with zeros. In cases when the support of the solution was well contained within the computational domain, this assumption did not lead to large numerical errors.

We note that in order to calculate (3.13), it would require $O(M^9)$ operations to calculate it, $O(M^6)$ at one velocity node for $O(M^3)$ velocity nodes. However, recall that the property stated in section 3.3 that only has $O(M^5)$ components bringing down the total cost of computing (3.13) down to $O(M^8)$ operations. However, we wish to reduce this further. It turns out that the Discrete Fourier Transform has properties that allow us to bring down the cost of computing (3.13) from $O(M^8)$ to $O(M^6)$.

Chapter 4

The Discrete Fourier Transform

The goal is to compute (3.13) faster than $O(M^6)$. We note that (3.13) is similar to a discrete convolution form. The Discrete Fourier Transform (DFT) can be used to speed up the computations of convolutions using the *Convolution Theorem*. In particular, it is used to speed up circular convolutions. In order to apply the DFT to our problem, we must first assume a periodic extension of the discrete form of f. This section discusses the background of the Discrete Fourier Transform, convolution theorem, and periodic continuation needed to quickly compute (3.13)

4.1 The One-Dimensional Discrete Fourier Transform and its Properties

The Discrete Fourier Transform is the discrete analog of the Fourier transform. It is a tool often used describe the relationship between the time and frequency representation of discrete signals [9].

Definition 1 *Let $\{x_n\}_{n=0}^{N-1}$ be a sequence of N complex numbers. The DFT is defined as*

$$\mathcal{F}[x]_k = \sum_{l=0}^{N-1} W^{lk} x_l, \qquad \textit{where} \quad W = \mathrm{e}^{-\imath 2\pi/N}\,. \tag{4.1}$$

4.1.1 Properties of the DFT

Many properties of the DFT are a natural consequence of its defintion. In this subsection, we review the properties of the DFT that will be important for applying the DFT to 3.13. First, the DFT is invertible.

Definition 2 *Let $\{\mathcal{F}[x]_k\}_{k=0}^{N-1}$ be a sequence of N complex numbers such that it is the*

DFT of the complex valued sequence $\{x_l\}_{l=0}^{N-1}$*. Then the inverse DFT is defined as*

$$x_l = \frac{1}{N}\sum_{k=0}^{N-1} W^{-lk}\mathcal{F}[x]_k\,. \tag{4.2}$$

We can see that definition of the inverse will give back the original sequence. If we have a sequence $\{x_l\}_{l=0}^{N-1}$, then note that if we take the inverse of $\mathcal{F}[x]_k$

$$\begin{aligned}\mathcal{F}^{-1}[\mathcal{F}[x]]_l &= \frac{1}{N}\sum_{k=0}^{N-1} W^{-lk}(\sum_{j=0}^{N-1} W^{kj}x_j)\\ &= \frac{1}{N}\sum_{k=0}^{N-1}\sum_{j=0}^{N-1} W^{-lk}W^{kj}x_j\\ &= \frac{1}{N}\sum_{j=0}^{N-1} x_j(\sum_{k=0}^{N-1} W^{(j-l)k})\\ &= N\frac{1}{N}x_l = x_l\end{aligned}$$

where the last step used that if $l \neq j$ then $\sum_{k=0}^{N-1} W^{(j-l)k} = 0$ and for $l = j$ we have that $\sum_{k=0}^{N-1} W^{(j-l)k} = \sum_{k=0}^{N-1} 1 = N$. Thus the definition given above returns the original sequence. To see this:

$$\begin{aligned}\mathcal{F}[z]_k &= \sum_{l=0}^{N-1} W^{kn}\sum_{n=0}^{N-1} x_n y_{l-n}\\ &= \sum_{n=0}^{N-1} x_n(\sum_{l=0}^{N-1} W^{kn}y_{l-n}\end{aligned}$$

The DFT is linear. That is, if we have two complex sequences $\{x_n\}_{n=0}^{N-1}$ and $\{y_n\}_{n=0}^{N-1}$, then

$$\mathcal{F}[x+y]_k = \mathcal{F}[x]_k + \mathcal{F}[y]_k\,. \tag{4.3}$$

This property will be used in order to apply the DFT to discrete convolution form of the collision integral.

The last property we need is the shift property. This says that if $\{y_n\}$ is a periodic sequence created by shifting $\{x_n\}$ by l, that is $\{y_n\} = \{x_{n+l}\}$, then

$$\mathcal{F}[y]_k = W^{lk}\mathcal{F}[x]_k \, . \tag{4.4}$$

4.1.2 The Convolution Theorem

Another important result is the convolution theorem. The result is stated and proved here. First we start with a definition.

Definition 3 *Let x_n and y_n be periodic sequences with period N. An N-point circular convolution of x_n and y_n is defined as (see, e.g., [9])*

$$z_l = \sum_{n=0}^{N-1} x_n y_{l-n} \, . \tag{4.5}$$

A well known property of the Fourier transform is that it converts circular convolution (4.5) into a product, namely,

$$\mathcal{F}[z]_k = \mathcal{F}[x]_k \mathcal{F}[y]_k \, . \tag{4.6}$$

After calculating $\mathcal{F}[z]_k$ using (4.6), the sequence $\{z_n\}_{n=0}^{N-1}$ can be retrieved via application of the inverse DFT, (4.2), e.g. $\mathcal{F}^{-1}[\mathcal{F}[z]]_n = z_n$.

The convolution theorem seemingly gives no computational gain. Straightforward computation of the DFT and inverse DFT will take $O(N^2)$ complex operations. Computation of (4.5) also takes $O(N^2)$ operations without needing to perform any complex multiplications if both $\{x_n\}$ and $\{y_n\}$ are real. However, the complexity of calculating (4.1) can be reduced by an algorithm known as the Fast Fourier Transform (FFT) which can calculate both the forward and inverse DFT in $O(N \log N)$ operations.

4.2 The One-Dimensional Fast Fourier Transform

In this section, technical details about the implementation of the Radix-2 Fast Fourier Transform are given. However, it should be noted that the Fast Fourier Transform has been implemented for many different radixes.

Chapter 5

Discretization of the Collision Integral and Fast Evaluation of Discrete Convolution

5.1 Formulas for Computing the DFT of the Collision Integral Copy and Pasted

To derive the formula for computing the collision operator we rewrite (3.13) as

$$I_{i;j_u,j_v,j_w} = \sum_{i',i''=1}^{s} I_{i,i',i'';j_u,j_v,j_w} \,, \tag{5.1}$$

where

$$I_{i,i',i'';j_u,j_v,j_w} = \sum_{j'_u,j'_v,j'_w=0}^{M-1} \sum_{j''_u,j''_v,j''_w=0}^{M-1} f_{i';j'_u-j_u,j'_v-j_v,j'_w-j_w} f_{i'';j''_u-j_u,j''_v-j_v,j''_w-j_w} A_{i',i'';j'_u,j'_v,j'_w,j''_u,j''_v,j''_w;i} \,.$$

In view of (5.1), we can focus on evaluation of $I_{i,i',i'';j_u,j_v,j_w}$. To simplify the notations in the discussion below, we drop the i, i', and i'' subscripts from $I_{i,i',i'';j_u,j_v,j_w}$, $f_{i';j'_u,j'_v,j'_w}$, $A_{i',i'';j'_u,j'_v,j'_w,j''_u,j''_v,j''_w;i}$, and $\mathcal{F}[I_{i,i',i''}]_{k_u,k_v,k_w}$ and write I_{j_u,j_v,j_w}, $f_{j'_u,j'_v,j'_w}$, $A_{j'_u,j'_v,j'_w,j''_u,j''_v,j''_w}$, and $\mathcal{F}[I]_{k_u,k_v,k_w}$, respectively. In particular, we have

$$I_{j_u,j_v,j_w} = \sum_{j'_u,j'_v,j'_w=0}^{M-1} \sum_{j''_u,j''_v,j''_w=0}^{M-1} f_{j'_u-j_u,j'_v-j_v,j'_w-j_w} f_{j''_u-j_u,j''_v-j_v,j''_w-j_w} A_{j'_u,j'_v,j'_w,j''_u,j''_v,j''_w} \,. \tag{5.2}$$

As is seen from definition (**??**), the multi-dimensional DFT results from applying the one-dimensional DFT along each dimension of the sequence for every fixed value of indices in the other dimensions (see e.g., [9]).

We fix indices j_v and j_w in equation (5.2) and apply the one-dimensional DFT in the remaining index j_u. Using linearity of the DFT and reordering the sums, we have

$$\mathcal{F}[I_{j_v,j_w}]_{k_u} = \sum_{j'_v,j'_w=0}^{M-1} \sum_{j''_v,j''_w=0}^{M-1} \mathcal{F}[\hat{I}_{j_v,j'_v,j''_v,j_w,j'_w,j''_w}]_{k_u} \, ,$$

where

$$\mathcal{F}[\hat{I}_{j_v,j'_v,j''_v,j_w,j'_w,j''_w}]_{k_u} = \sum_{j_u=0}^{M-1} W^{k_u j_u} \hat{I}_{j_u;j_v,j'_v,j''_v,j_w,j'_w,j''_w}$$

$$\hat{I}_{j_u;j_v,j'_v,j''_v,j_w,j'_w,j''_w} = \sum_{j''_u=0}^{M-1} \sum_{j'_u=0}^{M-1} f_{j'_u-j_u,j'_v-j_v,j'_w-j_w} f_{j''_u-j_u,j''_v-j_v,j''_w-j_w} A_{j'_u,j'_v,j'_w,j''_u,j''_v,j''_w} \, .$$

Once again, for purposes of calculating the one-dimensional discrete Fourier transform along dimension j_u, we need only consider the transform of $\hat{I}_{j_u,j_v,j'_v,j''_v,j_w,j'_w,j''_w}$. By similar argument as before, we fix and drop indices j_v, j'_v, j''_v, j_w, j'_w, and j''_w in the latter formula and write

$$\hat{I}_{j_u} = \sum_{j''_u=0}^{M-1} \sum_{j'_u=0}^{M-1} f_{j'_u-j_u} f_{j''_u-j_u,} A_{j'_u,j''_u} \tag{5.3}$$

$$\mathcal{F}[\hat{I}]_{k_u} = \sum_{j_u=0}^{M-1} \sum_{j''_u=0}^{M-1} \sum_{j'_u=0}^{M-1} W^{k_u j_u} f_{j'_u-j_u} f_{j''_u-j_u,} A_{j'_u,j''_u} \, . \tag{5.4}$$

We note that evaluating $\mathcal{F}[\hat{I}]_{k_u}$ directly would require $O(M^3)$ operations. However, taking into consideration the discussion in the last section, expression in the right side of (5.3) can be considered as a circular convolution, having a form similar to (**??**). This motivates us to explore properties of the DFT and rewrite (5.4) in a form suitable for numerical computation. This is accomplished in the following lemma.

Lemma 3 *Let $\{f_j\}_{j=0}^{M-1}$ be a M periodic sequence and $\{A_{ij}\}_{ij}$ be a two index sequence that is M periodic in both its indices. Let $\{\hat{I}_j\}_{j=0}^{M-1}$ be a new sequence defined by*

$$\hat{I}_j = \sum_{j'=0}^{M-1} \sum_{j''=0}^{M-1} f_{j'-j} f_{j''-j} A_{j',j''} \tag{5.5}$$

Let $\mathcal{F}[\hat{I}]_k$ be the DFT of $\hat{I}_j$, then

$$\mathcal{F}[\hat{I}]_k = M \sum_{l=0}^{M-1} \mathcal{F}^{-1}[f]_{k-l} \mathcal{F}^{-1}[f]_l \mathcal{F}[A]_{k-l,l} \tag{5.6}$$

Proof. Applying the one-dimensional DFT to $\hat{I}_j$, we have

$$\mathcal{F}[\hat{I}]_k = \sum_{j_u=0}^{M-1} \sum_{j'=0}^{M-1} \sum_{j''=0}^{M-1} W^{kj} f_{j'-j} f_{j''-j} A_{j',j''} \tag{5.7}$$

We define $\mathcal{F}[A_{j'}]_l$ to be the one-dimensional DFT of $A_{j',j''}$ in the second index, i.e.,

$$\mathcal{F}[A_{j'}]_l = \sum_{j''=0}^{M-1} W^{j''l} A_{j',j''} \,,$$

and rewrite $A_{j',j''}$ as

$$A_{j',j''} = \frac{1}{M} \sum_{l=0}^{M-1} W^{-j''l} \mathcal{F}[A_{j'}]_l \,. \tag{5.8}$$

Substituting (5.8) into (5.7), we have

$$\begin{aligned}
\mathcal{F}[\hat{I}]_k &= \frac{1}{M} \sum_{j=0}^{M-1} \sum_{j'=0}^{M-1} \sum_{j''=0}^{M-1} W^{jk} f_{j'-j} f_{j''-j} \left(\sum_{l=0}^{N-1} W^{-j''l} \mathcal{F}[A_{j'}]_l \right) \\
&= \frac{1}{M} \sum_{l=0}^{M-1} \sum_{j=0}^{M-1} \sum_{j'=0}^{M-1} \sum_{j''=0}^{M-1} W^{jk} f_{j'-j} f_{j''-j} W^{-j''l} \mathcal{F}[A_{j'}]_l
\end{aligned} \tag{5.9}$$

Consider the sum that runs over index j''. Assuming that indices l, j, and j' are held constant, we split the sum into two parts.

$$\begin{aligned}
&\sum_{j''=0}^{M-1} W^{jk} f_{j'-j} f_{j''-j} W^{-j''l} \mathcal{F}[A_{j'}]_l \\
&= \sum_{j''=0}^{j-1} W^{jk} f_{j'-j} f_{j''-j} W^{-j''l} \mathcal{F}[A_{j'}]_l + \sum_{j''=j}^{M-1} W^{jk} f_{j'-j} f_{j''-j} W^{-j''l} \mathcal{F}[A_{j'}]_l \,.
\end{aligned} \tag{5.10}$$

Notice $j'' - j < 0$ in the first sum. Using the periodicity of f_j the summation index can be redefined so that only values $f_{j''-j}$ with positive $j''-j$ appear in the sum. Indeed, we assume that $j'' < j$ and notice that $f_{j''-j} = f_{j''-j+M}$ since f_j is M periodic. We also have $W^M = 1$, so $W^{j''l} = W^{j''l+Ml} = W^{(j''+M)l}$. Introducing $\hat{j}'' = j'' + M$, we observe

$$\begin{aligned}\sum_{j''=0}^{j-1} W^{jk} f_{j'-j} f_{j''-j} W^{-j''l} \mathcal{F}[A_{j'}]_l &= \sum_{j''=0}^{j-1} W^{jk} f_{j'-j} f_{j''-j+M} W^{(-j''+M)l} \mathcal{F}[A_{j'}]_l \\ &= \sum_{\hat{j}''=M}^{M-1+j} W^{jk} f_{j'-j} f_{\hat{j}''-j} W^{-\hat{j}''l} \mathcal{F}[A_{j'}]_l \,.\end{aligned}$$

Combining the last formula with (5.10) we have

$$\sum_{j''=0}^{M-1} W^{jk} f_{j'-j} f_{j''-j} W^{-j''l} \mathcal{F}[A_{j'}]_l = \sum_{j''=j}^{M-1+j} W^{jk} f_{j'-j} f_{j''-j} W^{-j''l} \mathcal{F}[A_{j'}]_l \,. \tag{5.11}$$

Introducing a substitution of index $u'' = j'' - j$ we rewrite the right side of (5.11) as follows

$$\sum_{j''=0}^{M-1} W^{jk} f_{j'-j} f_{j''-j} W^{-j''l} \mathcal{F}[A_{j'}]_l = \sum_{u''=0}^{M-1} W^{jk} f_{j'-j} f_{u''} W^{(-u''-j)l} \mathcal{F}[A_{j'}]_l \,.$$

Going back to (5.9), we replace the inside sum with the last expression to have

$$\begin{aligned}&\frac{1}{M} \sum_{l=0}^{M-1} \sum_{j=0}^{M-1} \sum_{j'=0}^{M-1} \sum_{u''=0}^{M-1} W^{jk} f_{j'-j} f_{u''} W^{(-u''-j)l} \mathcal{F}[A_{j'}]_l \\ &\qquad = \frac{1}{M} \sum_{l=0}^{M-1} \sum_{j'=0}^{M-1} \sum_{u''=0}^{M-1} W^{-u''l} f_{u''} \left(\sum_{j=0}^{M-1} W^{j(k-l)} f_{j'-j} \mathcal{F}[A_{j'}]_l \right) \,.\end{aligned} \tag{5.12}$$

Now we focus on the term within the parentheses in (5.12). Splitting the sum and using periodicity, we obtain

$$\sum_{j=0}^{M-1} W^{j(k-l)} f_{j'-j}\mathcal{F}[A_{j'}]_l \tag{5.13}$$

$$\begin{aligned}
&= \sum_{j=0}^{j'} W^{j(k-l)} f_{j'-j}\mathcal{F}[A_{j'}]_l + \sum_{j=j'+1}^{M-1} W^{j(k-l)} f_{j'-j}\mathcal{F}[A_{j'}]_l \\
&= \sum_{j=0}^{j'} W^{j(k-l)} f_{j'-j}\mathcal{F}[A_{j'}]_l + \sum_{j=j'+1}^{M-1} W^{(j-M)(k-l)} f_{j'-j+M}\mathcal{F}[A_{j'}]_l \\
&= \sum_{j=0}^{j'} W^{j(k-l)} f_{j'-j}\mathcal{F}[A_{j'}]_l + \sum_{\hat{j}=j'-M+1}^{-1} W^{\hat{j}(k-l)} f_{j'-\hat{j}}\mathcal{F}[A_{j'}]_l \\
&= \sum_{j=j'-M+1}^{j'} W^{j(k-l)} f_{j'-j}\mathcal{F}[A_{j'}]_l = \sum_{u'=0}^{M-1} W^{(j'-u')(k-l)} f_{u'}\mathcal{F}[A_{j'}]_l\,.
\end{aligned} \tag{5.14}$$

Here $u' = j' - j$. Substituting this result into (5.12) and regrouping sums, we yield

$$\begin{aligned}
&\frac{1}{M}\sum_{l=0}^{M-1}\sum_{j'=0}^{M-1}\sum_{u''=0}^{M-1} W^{-u''l} f_{u''}\left(\sum_{j=0}^{M-1} W^{j(k-l)} f_{j'-j}\mathcal{F}[A_{j'}]_l\right) \\
&\quad = \frac{1}{M}\sum_{l=0}^{M-1}\sum_{j'=0}^{M-1}\sum_{u''=0}^{M-1} W^{-u''l} f_{u''}\left(\sum_{u'=0}^{M-1} W^{(j'-u')(k-l)} f_{u'}\mathcal{F}[A_{j'}]_l\right) \\
&\quad = M\sum_{l=0}^{M-1}\left(\frac{1}{M}\sum_{u'=0}^{M-1} W^{-u'(k-l)} f_{u'}\right)\left(\frac{1}{M}\sum_{u''=0}^{M-1} W^{-u''l} f_{u''}\right)\left(\sum_{j'=0}^{M-1} W^{j'(k-l)}\mathcal{F}[A_{j'}]_l\right).
\end{aligned} \tag{5.15}$$

The terms in the parentheses in (5.15) are just the definitions of the DFT. Thus we can write the equation as

$$\mathcal{F}[\hat{I}]_k = M\sum_{l=0}^{M-1} \mathcal{F}^{-1}[f]_{k-l}\mathcal{F}^{-1}[f]_l\mathcal{F}[A]_{k-l,l}\,.$$

■

Lemma 3 allows us to compute (5.4) in $O(M^2)$ operations. Indeed, it takes $O(M \log M)$ operations to compute $\mathcal{F}^{-1}[f]_{k_u}$ using a fast Fourier transform and it takes $O(M^2)$ operations to compute discrete convolution in the frequency space (5.6). To extend this result to $\mathcal{F}[I]_{k_u,k_v,k_w}$, it is sufficient to repeat the approach for indices j_v and j_w focusing on one dimension at a time. The following theorem summarizes the result.

Theorem 1 *Let f_{j_u,j_v,j_w} be a three-index sequence that is periodic in each index with period M and let $A_{j'_u,j'_v,j'_w,j''_u,j''_v,j''_w}$ be a M-periodic six-dimensional tensor. The multi-dimensional discrete Fourier transform of equation (5.2) can be represented as*

$$\mathcal{F}[I]_{k_u,k_v,k_w} = M^3 \sum_{l_u,l_v,l_w=0}^{M-1} \mathcal{F}^{-1}[f]_{k_u-l_u,k_v-l_v,k_w-l_w} \mathcal{F}^{-1}[f]_{l_u,l_v,l_w} \mathcal{F}[A]_{k_u-l_u,k_v-l_v,k_w-l_w,l_u,l_w,l_w} \tag{5.16}$$

Proof. We apply the one dimensional discrete Fourier transform along j_u in equation (5.2) and apply Lemma 3:

$$\begin{aligned}
\mathcal{F}[I_{j_v,j_w}]_{k_u} &= \sum_{j'_v,j'_w=0}^{M-1} \sum_{j''_v,j''_w=0}^{M-1} \left(\sum_{j_u=0}^{M-1} \sum_{j'_u=0}^{M-1} \sum_{j''_u=0}^{M-1} W^{j_u k} f_{j'_u-j_u,j'_v-j_v,j'_w-j_w} f_{j''_u-j_u,j''_v-j_v,j''_w-j_w} A_{j'_u,j'_v,j'_w,j''_u,j''_v,j''_w} \right) \\
&= M \sum_{j'_v,j'_w=0}^{M-1} \sum_{j''_v,j''_w=0}^{M-1} \sum_{l_u=0}^{M-1} \mathcal{F}^{-1}[f_{j'_v-j_v,j'_w-j_w}]_{k_u-l_u} \mathcal{F}^{-1}[f_{j''_v-j_v,j''_w-j_w}]_{l_u} \mathcal{F}[A_{j'_v,j'_w,j''_v,j''_w}]_{k_u-l_u,l_u} \\
&= M \sum_{l_u=0}^{M-1} \sum_{j'_w=0}^{M-1} \sum_{j''_w=0}^{M-1} \\
&\qquad \left(\sum_{j'_v=0}^{M-1} \sum_{j''_v=0}^{M-1} \mathcal{F}^{-1}[f_{j'_v-j_v,j'_w-j_w}]_{k_u-l_u} \mathcal{F}^{-1}[f_{j''_v-j_v,j''_w-j_w}]_{l_u} \mathcal{F}[A_{j'_v,j'_w,j''_v,j''_w}]_{k_u-l_u,l_u} \right).
\end{aligned} \tag{5.17}$$

We now focus on the terms inside the parentheses. We fix the indices j_w, j'_w, j''_w,

k_u, and l_u in the grouped terms We drop these indices and write

$$\tilde{I}_{j_v} = \sum_{j'_v=0}^{M-1} \sum_{j''_v=0}^{M-1} \tilde{f}_{j'_v - j_v} \tilde{f}_{j''_v - j_v} \tilde{A}_{j'_v, j''_v} \, ,$$

where

$$\tilde{f}_{j'_v - j_v} = \mathcal{F}^{-1}[f_{j'_v - j_v}], \qquad \tilde{A}_{j'_v, j''_v} = \mathcal{F}[A_{j'_v, j''_v}].$$

We can see that this expression is identical to (5.5). We take the discrete Fourier transform along the j_v index of $\mathcal{F}[\tilde{I}]_{k_v}$ and apply Lemma 3 to arrive at

$$\mathcal{F}[\tilde{I}]_{k_v} = M \sum_{l_v=0}^{M-1} \mathcal{F}^{-1}[\tilde{f}]_{k_v - l_v} \mathcal{F}^{-1}[\tilde{f}]_{l_v} \mathcal{F}[\tilde{A}]_{k_v - l_v, l_v} \, . \tag{5.18}$$

We recall definitions of $\tilde{f}_{j'_v - j_v}$ and $\tilde{A}_{j'_v, j''_v}$ and notice that the multi-index Fourier transform results from applying the one-dimensional transform in each index. Bringing indices j'_w, j''_w, k_u and l_u back, equation (5.18) becomes

$$\mathcal{F}[I_{j_w}]_{k_u, k_v} = M^2 \sum_{l_u, l_v=0}^{M-1} \mathcal{F}^{-1}[f_{j'_w - j_w}]_{k_u - l_u, k_v - l_v} \mathcal{F}^{-1}[f_{j''_w - j_w}]_{l_u, l_v} \mathcal{F}[A_{j'_w, j''_w}]_{k_u - l_u, k_w - l_w, l_u, l_w} \, .$$

Performing the discrete Fourier transform in the j_w and repeating the argument once more we arrive at the statement of the theorem.

■

5.2 The Algorithm and its Complexity Copy and Pasted

Theorem 1 allows us to calculate the collision operator (**??**) in $O(s^3 M^6)$ operations using the algorithm outlined below. We note that $\mathcal{F}[A_{i,i',i''}]_{k_u,k_v,k_w,l_u,l_w,l_w}$ can be precomputed and therefore does not factor into the algorithmic complexity analysis.

1. The first step of the algorithm is to evaluate $\mathcal{F}^{-1}[f_i]_{k_u,k_v,k_w}$. Evaluation of the

inverse Fourier transform requires $O(M^3 \log M)$ operations for each value of index i by utilizing three-dimensional FFT. This must be repeated for each i, resulting in the total of $O(s^3 M^3 \log M)$ operations where s^3 is the number of velocity nodes in each velocity cell.

2. Next we directly compute the convolution

$$\mathcal{F}[I_{i,i',i''}]_{k_u,k_v,k_w} = M^3 \sum_{l_u,l_v,l_w=0}^{M-1} \mathcal{F}^{-1}[f_{i'}]_{k_u-l_u,k_v-l_v,k_w-l_w} \mathcal{F}^{-1}[f_{i''}]_{l_u,l_v,l_w} \mathcal{F}[A_{i,i',i''}]_{k_u-l_u,k_v-l_v,k_w-l_w,l_u,l}$$

   using periodicity of both $\mathcal{F}^{-1}[f_i]_{l_u,l_v,l_w}$ and $\mathcal{F}[A_{i,i',i''}]_{k_u,k_v,k_w,l_u,l_w,l_w}$.

   For fixed values of indices i, i', i'' and k_u, k_v, k_w, calculating $\mathcal{F}[I_{i,i',i''}]_{k_u,k_v,k_w}$ requires $O(M^3)$ arithmetic operations. There are M^3 combinations of k_u, k_v, k_w and s^3 combinations of indices i, i', and i'', therefore complexity of this step is $O(s^3 M^6)$.

3. Linearity of the Fourier transform allows us to sum $\mathcal{F}[I_{i,i',i''}]_{k_u,k_v,k_w}$ along i', i'' to calculate $\mathcal{F}[I_i]_{k_u,k_v,k_w}$.

$$\mathcal{F}[I_i]_{k_u,k_v,k_w} = \sum_{i',i''=1}^{s} \mathcal{F}[I_{i,i',i''}]_{k_u,k_v,k_w} \,.$$

   This step requires adding s^2 sequences of length M^3 for every value of i, resulting in a complexity of $O(s^3 M^3)$ operations.

4. We recover $\mathcal{F}^{-1}[\mathcal{F}[I_i]]_{j_u,j_v,j_w} = I'_{i;j_u,j_v,j_w}$. This requires calculating the three-dimensional inverse DFT for every i which gives a complexity of $O(sM^3 \log M)$.

Overall, the algorithm has the numerical complexity of $O(s^3 M^6)$ dominated by step 2. We note that $s = s_u s_v s_w$ is usually kept fixed and the number of cells M^3 in velocity domain is changing. In this case, the main contribution to complexity growth comes from M, the number of velocity cells in one velocity dimension. Thus we can consider the algorithm to be of complexity $O(M^6)$. In our simulations $s_u, s_v, s_w \leq 3$, however higher values may be used too, at least theoretically. The results of this analysis are validated within the next chapter.

needs rewriting In the implementation of the algorithm, we use the tensor product ordering for f. However, after taking the inverse FFT of $\mathcal{F}[I_i]_{i;j_u,j_v,j_w}$, we get differently ordered sequence, $I'_{i;j_u,j_v,j_w}$. This is due to the fact we take the FFT with respect to j, the shift in velocity cell. To retrieve back $I_{i;j_u,j_v,j_w}$ from $I'_{i;j_u,j_v,j_w}$, the following equation was used:

$$j_u = (j'_u - j_{u;c} \mod M)\,,\ j_v = (j'_v - j_{v;c} \mod M)\,,\ j_w = (j'_w - j_{w;c} \mod M)\,,$$

where $j_u u; c$ is canonical node, j'_u is the old index in $I_{i;j_u,j_v,j_w}$, and the $\mod$ operation always returns the positive remainder. To be explicit, the algorithm below

Algorithm 1 Perform FFT convolution at a single point in space.

```
function IndicesKeepTrack($f$,$\mathcal{F}[A]$)
    Input $f$, $Q$                                      ▷ The solution at time $t$.
    Input $\mathcal{F}[A]$                              ▷ The DFT of the $A$ operator.
    for i=1,s do
        $F_i = \mathcal{F}^{-1}[f]$        ▷ Calculate the inverse DFT of $f$ along velocity cells.
    end for
    for Each combination of $i, i', i'' = 1, s$ do
        $\mathcal{F}[I_{i,i',i''}] = M^3 \sum_{l_u,l_v,l_w=0}^{M-1} \mathcal{F}^{-1}[f_{i'}]_{k_u-l_u,k_v-l_v,k_w-l_w} \mathcal{F}^{-1}[f_{i''}]_{l_u,l_v,l_w} \mathcal{F}[A_{i,i',i''}]_{k_u-l_u,k_v-l_v,k_w-l_w,l_u,}$
        $\mathcal{F}[I_i] = \mathcal{F}[I_i] + \mathcal{F}[I_{i,i',i''}]$
    end for
    $I'_i = \mathcal{F}^{-1}[\mathcal{F}[I_i]]$        ▷ Take inverse DFT of $\mathcal{F} I_i$ along velocity cells.
    for i=1,s do                                        ▷ Reordering step.
        for j=1,M do
            $j'_u = j_u - j_{u;c} \mod M$
            $j'_v = j_u - j_{v;c} \mod M$
            $j'_w = j_w - j_{w;c} \mod M$
            $I_{j_u,j_v,j_w;i} = I_{j'_u,j'_v,j'_w;i}$
        end for
    end for
    return $I_i$
end function
```

5.3 Alternative Formulation Involving A^+ Discretization

Chapter 6

Numerical Results

This chapter will cover the numerical results of computing the collision operator using the nodal-DG velocity discretizations and convolution using the DFT.

6.1 Reduction in Computational Complexity

This section will cover estimating the numerical complexity of the method. The analysis in section 5.2 suggested that the algorithm had a total number of operations of $O(M^6)$ where M is the number of velocity cells in one velocity dimension. As discussed in [2], the direction evaluation of convolution (3.13) requires $O(M^8)$ operations. Thus we expect the FFT evaluation of (3.13) to take much less time than the method discussed in [2].

The results in Table 6.1 display the CPU times for evaluating the collision operator at one spatial point using both the FFT and direct evaluations. The computations were done on a Intel Core i7-3770 3.4 GHz processor. The code was compiled using the Intel Fortran Compilers along with Intel Math Kernal Library. The number of cells in the velocity domain were varied from 9 to 27. $s = 1$ in each run. The computational complexity was modeled using $t = O(M^\alpha)$ where α is a constant. The estimated values of α were calculated as $\hat{\alpha} = \ln(M_1/M_2)/\ln(t_1/t_2)$.

Copy and Pasted Paragraph Note that in the case of the Fourier evaluation, the observed orders are significantly higher than the projected value of 6. Still, the orders are significantly lower than the orders of the direct evaluation. Deviations from the theoretical estimate of $\alpha = 6$ may be due to the costs of the memory transfer operations and due to the choice of the specific fast Fourier transform that was automatically selected by the KML library based on the value of M. Overall, the new approach showed a dramatic improvement in speed as compared to the direct evaluation of the collision operator used in [1]. The acceleration is expected to be even larger for higher values of M.

	FFT		Direct		Speedup
M	time, s	$\hat{\alpha}$	time, s	$\hat{\alpha}$	
9	1.47E-02		1.25E-01		8.5
15	3.94E-01	6.43	4.91E+00	7.18	12.5
21	3.09E+00	6.14	7.80E+01	8.21	25.2
27	1.64E+01	6.65	6.05E+02	8.15	36.7

Table 6.1: CPU times for evaluating the collision operator directly and using the Fourier transform.

6.2 The Micro-Macro Decomposition

The form of the collision integral can be reformulated for the purpose of numerical implementation. During simulations, it was observed that the ability of numerical simulations to conserve mass, momentum, and energy is strongly affected by the form of the discrete collision integral.

One such reformulation is the micro-macro decomposition [**??**] which considers the solution has the sum of the target Maxwellian distribution and the deviation from the Maxwellian, i.e.,

$$f(t,\vec{x},\vec{v}) = f_M(t,\vec{x},\vec{v}) + \Delta f(t,\vec{x},\vec{v})\,, \tag{6.1}$$

where $f_M(t,\vec{x},\vec{v})$ takes the form (2.14) with the same temperature, bulk velocity, and temperature as $f(t,\vec{x},\vec{v})$. This form was applied in [1] in order to improve conservation properties of the scheme when the solution is near continuum. The numerical implementation used for the FFT convolution also used this decomposition.

We substitute (6.1) into (3.7) to obtain an alternative form of the collision integral:

$$\begin{aligned}
I_{\phi_{i;j}} &= \int_{\mathbb{R}^3}\int_{\mathbb{R}^3} f(t,\vec{x},\vec{v})f(t,\vec{x},\vec{v}_1)A(\vec{v},\vec{v}_1;\phi_{i;j})d\vec{v}_1\,d\vec{v} \\
&= \int_{\mathbb{R}^3}\int_{\mathbb{R}^3} [f_M(t,\vec{x},\vec{v})\Delta f(t,\vec{x},\vec{v}_1) + \Delta f(t,\vec{x},\vec{v})f_M(t,\vec{x},\vec{v}_1) \\
&\qquad + \Delta f(t,\vec{x},\vec{v})\Delta f(t,\vec{x},\vec{v}_1)]A(\vec{v},\vec{v}_1;\phi_{i;j})d\vec{v}_1\,d\vec{v}\,. \qquad (6.2)
\end{aligned}$$

where we used that $I[f_M](t,\vec{x},\vec{v}) = 0$. It was found that this form of the operator provided much better numerical results. Thus the numerical results in the rest of this work use (6.2) for all the calculations.

6.3 Numerical Results of the Split Form of the Operator

As was discussed in section 3.3, there is an alternative formulation of (3.7) which we write as (3.9). This section discusses the numerical issues with conservation that this form had.

Begin copy and paste In Figure 6.1 results of the evaluation of the collision integral at a single spatial point are presented for both split and non-split formulations. The value of the solution $f(t,\vec{v})$ in these computations is given by the sum of two Maxwellian distributions with dimensionless densities, bulk velocities, and temperatures given as follows: $n_1 = 1.6094$, $n_2 = 2.8628$, $\bar{\vec{u}}_1 = (0.7750, 0, 0)$, $\bar{\vec{u}}_2 = (0.4357, 0, 0)$, $T_1 = 0.3$, and $T_2 = 0.464$. These values correspond to upstream and downstream conditions of a normal shock wave with the Mach number 1.55. Discretization of the solution was done using 27 velocity cells in each dimension and one velocity node per cell. The collision operator was evaluated using both the split and non-split forms and using both direct evaluation and evaluation using the Fourier transform. Results of the direct evaluation of the split form of the collision operator are shown in plots (b) and (e). Notably, values of the collision operator are zero at the boundary of the domain, which is what one would expect from the collision process. Results of evaluation of the split form of the collision operator using the Fourier transform are shown in plots (a) and (b). Significant non-zero values can be observed at the corners of the domain. This is likely to be a manifestation of aliasing. Results of evaluating the non-split form of the collision operator using the Fourier transform are shown in plots (c) and (f). One can notice that in the case of the non-split form, aliasing is not visible. In fact, the L^1-norm of the difference between the direct and Fourier evaluations of the non-split collision operator in this case was 2.9E-4 and the L^∞-norm was 1.1E-4. We note that the diameters of the support of the collision kernels are comparable in both split and non-split cases. However, the non-split kernel is more sparse. We also note that by padding the solution and the collision kernel with zeros, the aliasing errors can be reduced in the case of the split formulations. However, this will also increase memory and time costs of calculations. The non-split form of the collision operator has significantly smaller aliasing errors and does not require zero padding. Therefore, it is more efficient.

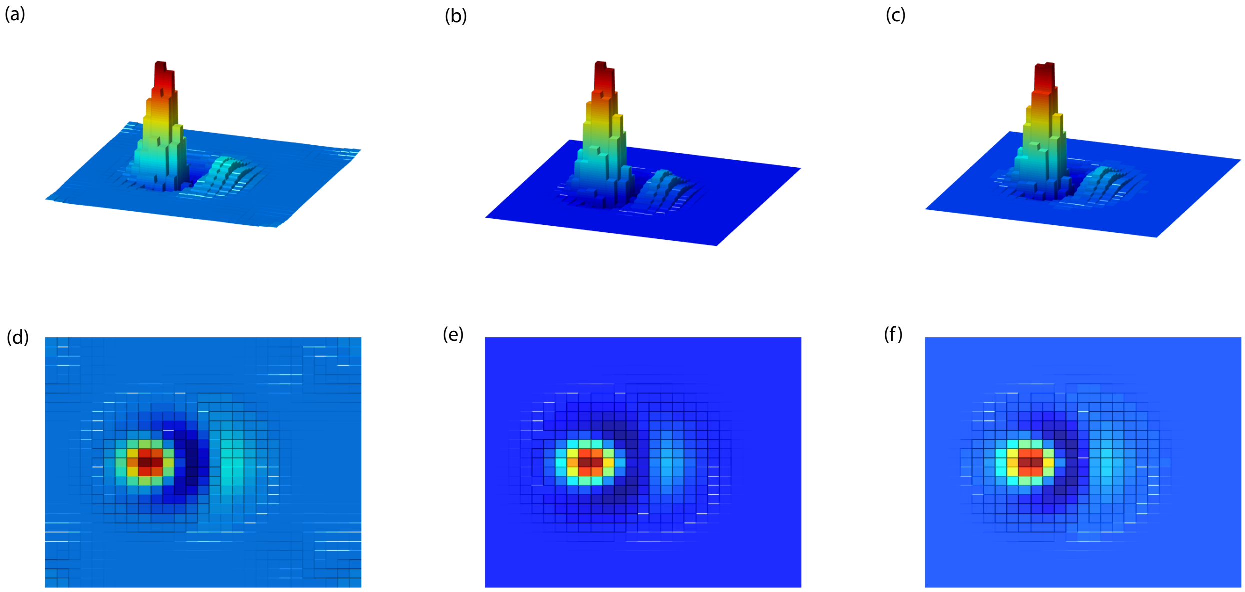

Figure 6.1: Evaluation of the collision operator using split and non-split forms: (a) and (d) the split form evaluated using the Fourier transform; (b) and (e) the split form evaluated directly; (c) and (f) the non-split form evaluated using the Fourier transform.

Another important issue that makes the non-split formulation more attractive is concerned with conservation of mass, momentum, and energy in the discrete solutions. It is the property of the exact Boltzmann collision operator that its mass, momentum, and temperature moments are zero. Generally, the conservation laws are satisfied only approximately when the Boltzmann equation is discretized. Many numerical approaches include mechanisms dedicated to enforcement of the conservation laws in discrete solutions in order to guarantee a physically meaningful result.

Error in Conservation of Mass					Error in Conservation of Temperature			
	Split		Non-split		Split		Non-split	
n	Fourier	Direct	Fourier	Direct	Fourier	Direct	Fourier	Direct
9	0.37	1.26	1.71E-5	1.92E-5	3.51	1.69	1.71E-2	1.84E-2
15	0.10	1.20	1.45E-5	1.71E-5	0.29	1.25	1.64E-3	3.15E-3
21	0.18	1.18	0.67E-5	0.93E-5	1.38	1.24	5.61E-5	1.75E-3
27	0.18	1.18	0.61E-5	0.86E-5	1.37	1.24	5.40E-4	1.05E-3

Table 6.2: Absolute errors in conservation of mass and temperature in the discrete collision integral computed using split and non-split formulations.

It was observed that if no measures are introduced to enforce the conservation laws, solutions to the problem of spatially homogeneous relaxation obtained using the split formulation of the collision integral exhibit large, on the order of 5% errors in temperature. The mass and momentum are also poorly conserved in this case. At the same time, solutions obtained using the non-split formulation had their mass, momentum, and temperature accurate to three or more digits. To further explore this phenomena, we evaluated the collision operator in both split and non-split forms and computed its mass, momentum, and temperature moments. The solution was taken to be the sum of two Maxwellians in the example above. The numbers of velocity cells were varied from 9 to 27. In both split and non-split formulations of the collision integral, the decomposed form (**??**) of the solution was used. For both forms, evaluation of the collision operator was done directly and using the Fourier transform. The results are summarized in Table 6.2. It can be seen that errors in the mass and temperature in the non-split formulation are several orders of magnitude smaller than in the split formulation. The errors are also larger in the case of direct evaluation. A possible explanation to this is the combined effect of finite precision arithmetic and truncation errors in integration that lead to catastrophic cancellation when gain and loss terms are combined. We note that in both split and non-split forms, fulfilment of conservation laws requires exact cancellation of the respective integration sums. When the gain and loss terms are computed separately using numerical quadratures, the relative truncation errors are expected to be acceptable for each of the terms. This may change, however, when the terms are combined. It is conceivable that significant digits cancel in the two terms and the truncation errors are promoted into significance, manifesting in strong violations of conservation laws. At the same time, increasing the number of velocity cells may not remedy the problem due to the expected accumulation of roundoff errors. Indeed, evaluation of the gain term in (3.9) requires $O(M^8)$ arithmetic operations. It is possible that combination of large and small values in the finite precision arithmetic results in loss of low order digits and a significant accumulation of roundoff. When the gain and loss terms are combined, this, again, will lead to loss of significance and to perturbations of conservation laws. In the case when both the non-split form and the decomposition (6.2) are used, much of the cancellation is happening on the level of the integrand. We hypothesize here that the resulting values of the integrand are smaller and vary less in scale. As a result, the accumulated absolute truncation and roundoff errors are also smaller, which gives better accuracy in conservation laws.

Because of the poor conservation properties and because of the susceptibility to aliasing errors we do not recommend the split form (3.9) for numerical implementation. **end copy and paste**

6.4 0d Homogeneuous Relaxation COPY AND PASTED

In this section we present results of solution of the problem of spatially homogeneous relaxation using Fourier evaluation of the collision operator. Two cases of initial data were considered. In both cases, the initial data is a sum of two Maxwellian densities. In the first case, the dimensionless densities, bulk velocities, and temperatures of the Maxwellians are $n_1 = 1.0007$, $n_2 = 2.9992$, $\bar{\vec{u}}_1 = (1.2247, 0, 0)$, $\bar{\vec{u}}_2 = (0.4082, 0, 0)$, $T_1 = 0.2$, $T_2 = 0.7333$. These parameters correspond to upstream and downstream conditions of the Mach 3 normal shock wave. In the second case, we use the parameters of the example of the previous section: $n_1 = 1.6094$, $n_2 = 2.8628$, $\bar{\vec{u}}_1 = (0.7750, 0, 0)$, $\bar{\vec{u}}_2 = (0.4357, 0, 0)$, $T_1 = 0.3$, and $T_2 = 0.464$. These parameters correspond to upstream and downstream conditions of a Mach 1.55 shock wave.

In Figures 6.2 and 6.3, relaxation of moments in the Mach 3.0 and Mach 1.55 solutions are presented. In the case of Mach 3.0, $M = 33$ velocity cells were used in each velocity dimension with one velocity node on each cell, $s = 1$. In the case of Mach 1.55, $M = 15$ and $s = 1$ were used. In the computed solutions, the collision operator was evaluated both using the Fourier transform and directly. In the Mach 3.0 instance, the directional temperature moments were compared to the moments obtained from a DSMC solution [**?**].

It can be seen that the solutions obtained by the Fourier evaluation of the collision integral are close to those computed by the direct evaluation. The low order moments are in excellent agreement for both presented solutions. However, there are differences

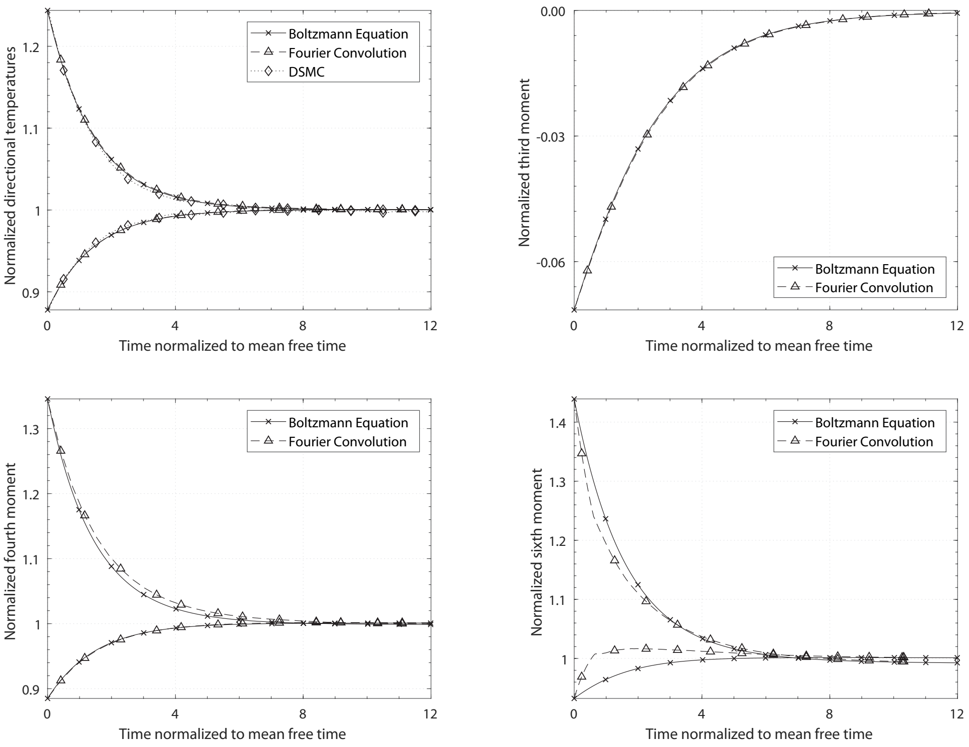

Figure 6.2: Relaxation of moments $f_{\varphi_{i,p}} = \int_{R^3}(u_i - \bar{u}_i)^p f(t,\vec{u})\,du$, $i = 1, 2$, $p = 2, 3, 4, 6$ *in a mix of Maxwellian streams corresponding to a shock wave with Mach number 3.0 obtained by solving the Boltzmann equation using Fourier and direct evaluations of the collision integral. In the case of* $p = 2$*, the relaxation of moments is also compared to moments of a DSMC solution [**?**].*

in the higher moments. It appears that the differences are caused by a small amount of the aliasing error in the solutions. This can be reduced by padding the solution and the kernel with zeros at the expense of higher numerical costs, both in time and memory. Overall, however, the $O(M^6)$ evaluation of the collision operator using the Fourier transform appears to be consistent and stable.

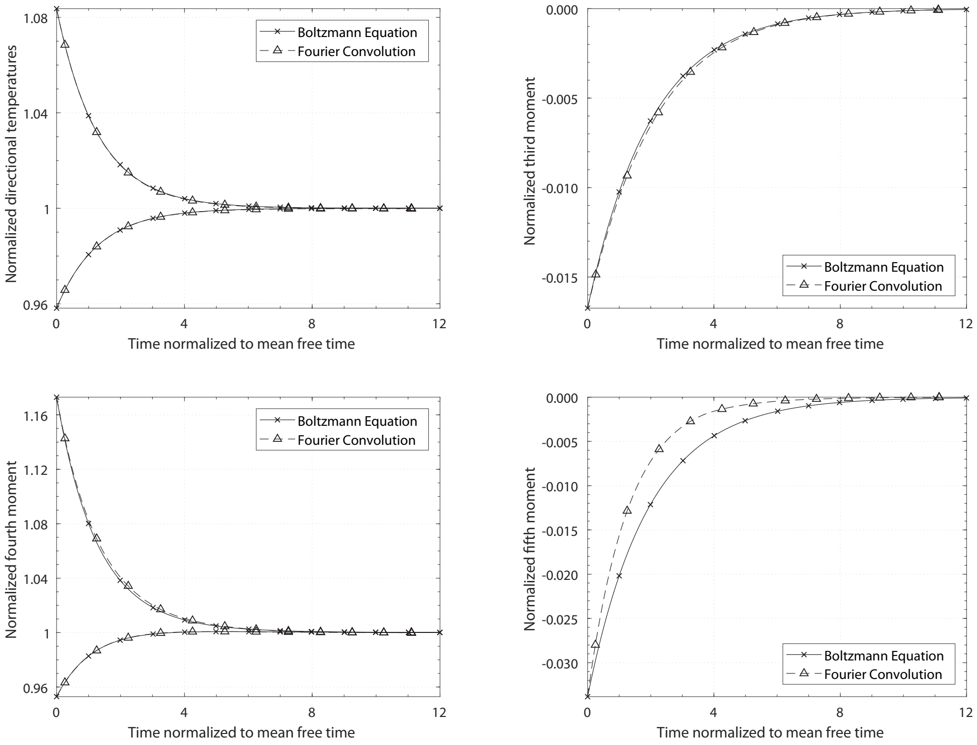

Figure 6.3: Relaxation of moments $f_{\varphi_{i,p}}$, $i = 1, 2$, $p = 2, 3, 4, 6$ *in a mix of Maxwellian streams corresponding to a shock wave with Mach number 1.55 obtained by solving the Boltzmann equation using Fourier and direct evaluations of the collision integral.*

References

[1] A. Alekseenko and E. Josyula. Deterministic solution of the Boltzmann equation using a discontinuous Galerkin velocity discretization. In *28th International Symposium on Rarefied Gas Dynamics, 9-13 July 2012, Zaragoza, Spain*, AIP Conference Proceedings, page 8. American Institute of Physics, 2012.

[2] A. Alekseenko and E. Josyula. Deterministic solution of the spatially homogeneous boltzmann equation using discontinuous galerkin discretizations in the velocity space. *Journal of Computational Physics*, 272(0):170 – 188, 2014.

[3] A. Alekseenko, T. Nguyen, and A. Wood. A deterministic-stochastic method for computing the boltzmann collision integral in $\mathcal{O}(mn)$ operations. *submitted to Kinetic and Related Models*, page 30p, 2015.

[4] L. Boltzmann. Weitere studien u ber das warmegleichgewicht unter gasmolekulen. 1874.

[5] J.S. Hesthaven and T. Warburton. *Nodal Discontinuous Galerkin Methods: Algorithms, Analysis, and Applications.* Texts in Applied Mathematics. Springer, 2007.

[6] M.N. Kogan. *Rarefied Gas Dynamics.* Plenum Press, New York, USA, 1969.

[7] G. Kremer. *An Introduction to the Boltzmann Equation and Transport Processes in Gases.* Springer, New York, 2010.

[8] Armando Majorana. A numerical model of the boltzmann equation related to the discontinuous Galerkin method. *Kinetic and Related Models*, 4(1):139 – 151, March 2011.

[9] H. J. Nussbaumer. *Fast Fourier Transform and Convolution Algorithms*. Springer Series in Information Sciences. Springer-Verlag, Heidelberg, 1982.

[10] H. Struchtrup. *Macroscopic Transport Equations for Rarefied Gas Flows Approximation Methods in Kinetic Theory*. Interaction of Mechanics and Mathematics Series. Springer, Heidelberg, 2005.